

# What is Linear Regression?



# Supervised Learning: Regression

**“Supervised”?**

- **A set of samples where the desired labels are known**

# **What is the main goal of supervised learning?**

- **Learn a model from labeled training data that allows for predictions on new unseen data**

# Supervised Learning: Regression

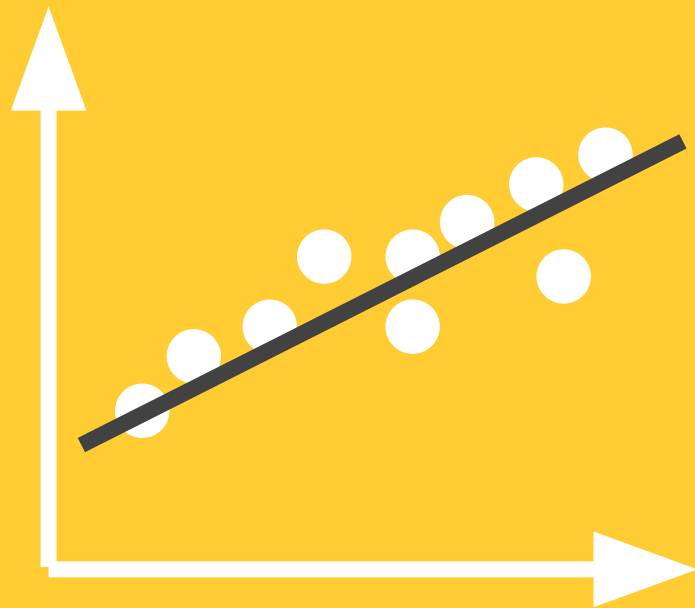
- **Given a training set of labeled input-output pairs**
  - $(x_1, y_1), (x_2, y_2) \dots (x_N, y_N)$
  - $x$  and  $y$  can be any value
  - each  $y$  is generated by an unknown function  $y = f(x)$

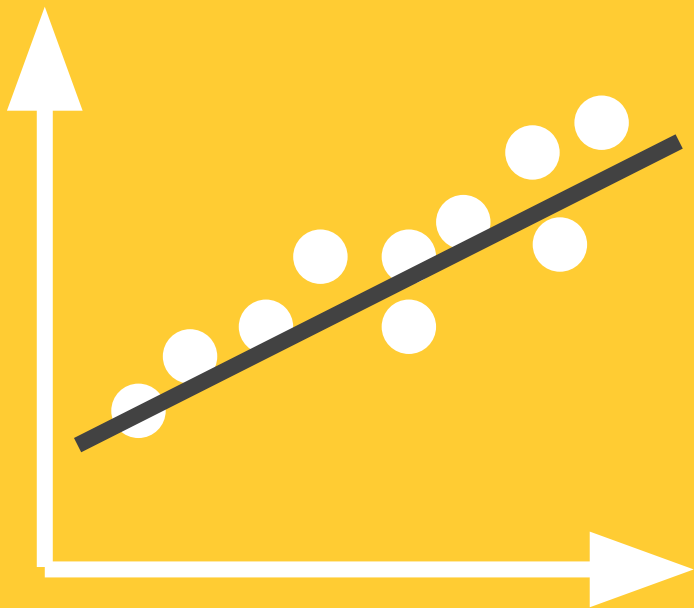
# Supervised Learning: Regression

- What is the goal?
  - Find a hypothesis  $h$  that approximates the true function  $f$

# What is regression?

- Subcategory of supervised learning
- What is the goal?
  - predict continuous outcomes





- Data generated by true function  $f$  + noise



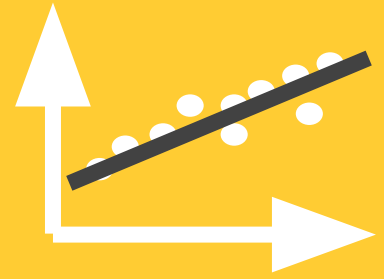
### Hypothesis $h$

- Equation  $y = mx + b$
- $M$  = slope
- $B$  = y-intercept

# Regression

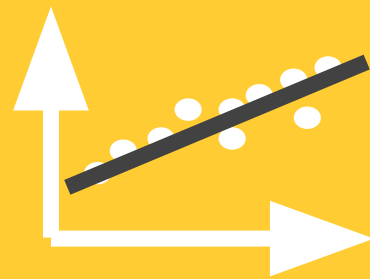
- Given a number of features and a continuous outcome, find a relationship between the features to predict an outcome



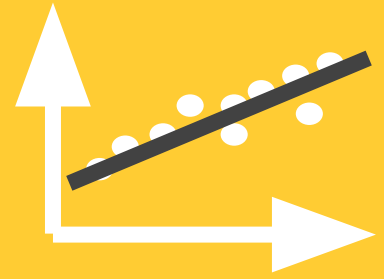


- Assumes the regression function is linear
  - Simple
- 
- Often adequate and interpretable description of how inputs affect outputs

- **Sometimes outperform non linear models!**



- **Especially with:**
  - small training datasets, or
  - low signal to noise ratio, or
  - sparse data

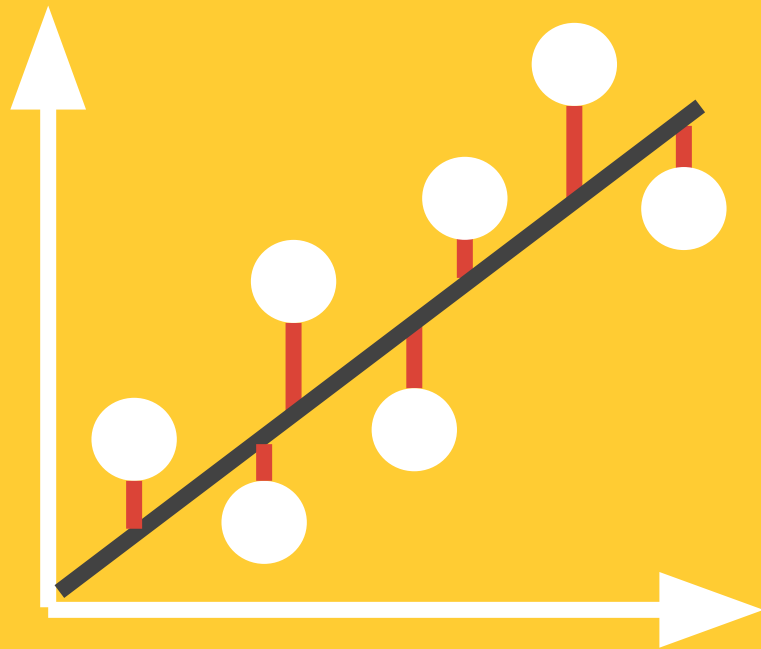


# Linear Model Fit By Least Squares Method

- Prediction method
- Mainstay of statistics
- One of the most important tools

# Linear Regression: Machine Learning

- What is the goal?
  - minimize **error**



# What is Linear Regression?

